

Amendments to the Specification:

Please replace paragraph [0025] with the following amended paragraph:

[0025] The second step 200 may be performed by hand or by a computer software program such as, for example, the program available from the University of Washington called "RepeatMasker", a program that recognizes sequences that are repeated in the genome (A. F. A. Smit and P. Green, www.genome.washington.edu/uwgc/analysistools/repeatmask, incorporated herein by reference). Essentially, RepeatMasker screens genomic sequences for repeat regions in DNA, referencing a database of known repetitive elements called RepBase. RepBase Version 5 has been employed in the methods of the present invention, as have earlier versions of RepBase. The RepBase database can be licensed from the Genetic Information Research Institute (see www.girinst.org, incorporated herein by reference). Essentially, known repetitive sequences such as Single Interspersed Nuclear Elements (SINEs, such as alu and MIR sequences), Long Interspersed Nuclear Elements (LINEs such as LINE1 and LINE2 sequences), Long Terminal Repeats (LTRs such as MaLRs, Retrov and MER4 sequences), Transposons, MER1 and MER2 sequences are "masked" or removed by the RepeatMasker program by substituting each specific nucleotide of the repeated regions (A, T, G or C) with an "N" or "X". In addition, xprimer (alces.med.umn.edu, Virtual Genome Center, incorporated herein by reference), a primer selection tool described below, can be used to identify simple, complex and internal repeats from a small database of repeats. Also, NCBI offers an Electronic PCR feature through its website (ncbi.nlm.nih.gov, incorporated herein by reference). The Electronic PCR program removes repetitive sequences from a non-repetitive marker set.